

Data Exploration and Cleaning:

Expected database structure:

1. Orders: Contains details of customer orders.
 - order_id, customer_id, order_date, total_amount, status
2. Customers: Contains details about customers.
 - customer_id, name, email, location, signup_date, birthdate
3. Products: Details of the products in the store.
 - product_id, product_name, category, price
4. Order_Items: Details of products in each order.
 - order_id, product_id, quantity
5. Customer Reviews: Customer feedback on purchased products.
 - review_id, order_id, product_id, rating, review_text, review_date

Table creation:

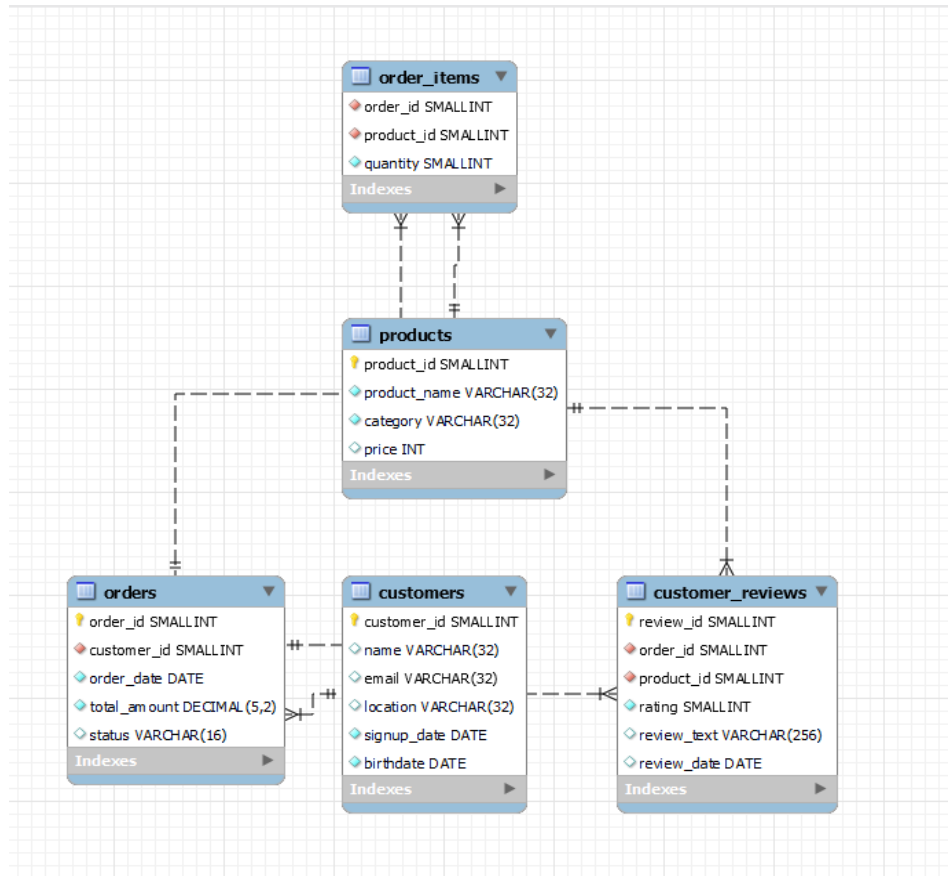
<pre>create table customers (customer_id smallint primary key, name varchar(32), email varchar(32), location varchar(32), signup_date date not null, birthdate date not null);</pre>	<pre>create table orders (order_id smallint primary key AUTO_INCREMENT, customer_id smallint not null, order_date date not null, total_amount decimal(5,2) not null, status varchar(16), foreign key (customer_id) references customers(customer_id));</pre>
<pre>create table products (product_id smallint primary key, product_name varchar(32) not null, category varchar(32) not null, price decimal(5,2) not null);</pre>	<pre>create table order_items (order_id smallint not null, product_id smallint not null, quantity smallint not null, foreign key(order_id) references orders(order_id), foreign key(product_id) references products(product_id)) ;</pre>
<pre>create table customer_reviews (review_id smallint primary key, order_id smallint not null, product_id smallint not null, rating smallint not null, review_text varchar(256), review_date date, foreign key(order_id) references orders(order_id), foreign key(product_id) references products(product_id));</pre>	

Database structure:

Info Tables Columns Indexes Triggers Views Stored Procedures Functions Grants Events													
Name	Engine	Version	Row Format	Rows	Avg Row Length	Data Length	Max Data Length	Index Length	Data Free	Auto Incre...	Create Time	Update Time	
customer_reviews	InnoDB	10	Dynamic	30	546	16.0 KB	0.0 bytes	32.0 KB	0.0 bytes	0	2025-03-31 16:21:36		
customers	InnoDB	10	Dynamic	11	1489	16.0 KB	0.0 bytes	0.0 bytes	0.0 bytes	0	2025-03-31 15:27:17		
order_items	InnoDB	10	Dynamic	37	442	16.0 KB	0.0 bytes	32.0 KB	0.0 bytes	0	2025-03-31 16:16:23		
orders	InnoDB	10	Dynamic	30	546	16.0 KB	0.0 bytes	16.0 KB	0.0 bytes	30	2025-03-31 15:28:04		
products	InnoDB	10	Dynamic	30	546	16.0 KB	0.0 bytes	0.0 bytes	0.0 bytes	0	2025-03-31 18:31:06		

Info Tables Columns Indexes Triggers Views Stored Procedures Functions Grants Events									
Table	Column	Type	Default Value	Nullable	Character Set	Collation	Privileges	Extra	Comments
customer_reviews	order_id	smallint		NO			select,insert,update,references		
customer_reviews	product_id	smallint		NO			select,insert,update,references		
customer_reviews	rating	smallint		NO			select,insert,update,references		
customer_reviews	review_date	date		YES			select,insert,update,references		
customer_reviews	review_id	smallint		NO			select,insert,update,references		
customer_reviews	review_text	varchar(256)		YES	utf8mb4	utf8mb4_0900_...	select,insert,update,references		
customers	birthdate	date		NO			select,insert,update,references		
customers	customer_id	smallint		NO			select,insert,update,references		
customers	email	varchar(32)		YES	utf8mb4	utf8mb4_0900_...	select,insert,update,references		
customers	location	varchar(32)		YES	utf8mb4	utf8mb4_0900_...	select,insert,update,references		
customers	name	varchar(32)		YES	utf8mb4	utf8mb4_0900_...	select,insert,update,references		
customers	signup_date	date		NO			select,insert,update,references		
order_items	order_id	smallint		NO			select,insert,update,references		
order_items	product_id	smallint		NO			select,insert,update,references		
order_items	quantity	smallint		NO			select,insert,update,references		
orders	customer_id	smallint		NO			select,insert,update,references		
orders	order_date	date		NO			select,insert,update,references		
orders	order_id	smallint		NO			select,insert,update,references	auto_increment	
orders	status	varchar(16)		YES	utf8mb4	utf8mb4_0900_...	select,insert,update,references		
orders	total_amount	decimal(5,2)		NO			select,insert,update,references		
products	category	varchar(32)		NO	utf8mb4	utf8mb4_0900_...	select,insert,update,references		
products	price	int		YES			select,insert,update,references		
products	product_id	smallint		NO			select,insert,update,references		
products	product_name	varchar(32)		NO	utf8mb4	utf8mb4_0900_...	select,insert,update,references		

ER diagram:



Data load:

Python with SQL connect Engine using sql alchemy.

```
[19]: import pandas as pd
import mysql.connector
from sqlalchemy import create_engine
```

```
[21]: engine = create_engine('mysql+mysqlconnector://root:newpass2024@127.0.0.1:3306/transactional_data')
```

```
[43]: data=pd.read_csv(r"E:\project_stuff\Data Analyst - Raw Data (1)\Senior Data Analyst - Raw Data\order_items.csv")
data.head()
```

```
[43]:
```

	order_id	product_id	quantity
0	1	1	2
1	1	3	1
2	2	4	1
3	3	6	1
4	3	8	1

```
[45]: data.to_sql('order_items',con=engine, if_exists='append', index=False)
```

```
[45]: 37
```